

. Starting Point for Enzymatic Hydrolysis for Cellulose: Enzyme
Engineering of Glycoside Hydrolase-5 Endoglucanases

Blake A. Simmons^{1,2}, Zhiwei Chen^{1,2}, Supratim Datta^{1,2}, Danielle T. Ercek¹, Hanbin Liu^{1,2}, Josh I. Park^{1,2}, Huu Tran^{1,2}, Swapnil R. Chhabra¹, Masood Z. Hadi^{1,2} and Rajat Sapra^{1,2} Joint BioEnergy Institute, Physical Biosciences Division, Lawrence Berkeley National Laboratory¹, and Sandia National Laboratories²

Enzymes are catalysts that have a set of specific and well defined characteristics like substrate, pH, temperature and kinetics etc. The vast amount of available genomic data shows that orthologous genes code for proteins that vary in these characteristics and the changes in properties at the protein level can be traced to sometimes minor changes like substitutions in the sequence at the molecular level. This observation that changes in the sequence leads to changes in characteristics can be used to design new and novel protein variants using two major approaches -directed evolution and rational design. We propose to use these techniques to engineer enzymes with improved activity and characteristics better suited to optimal conditions for cellulosic biomass deconstruction. Directed evolution of enzymes is a powerful technique that takes advantage of the Darwinian process of natural selection on a high throughput lab scale for the generation of protein mutants, called variants, which are then screened and selected for improved desirable traits as compared to the parent protein characteristics. At the core of the technique is the principle that incremental changes acquired either through mutagenesis or via recombination lead to a better variant when selection pressure is applied; the selection pressure can be any of the characteristics that are sought to be improved upon like activity, kinetics, pH or temperature stability and, in some cases, different substrates and novel reactions. The advantage for using the directed evolution approach is that there is no requirement for the availability of an extensive data set of orthologs as a starting point; only a gene sequence and a screening method for 'evolving' a protein function is needed. Therefore, this technique can be employed in all the enzymes currently known to be involved in cellulosic biomass deconstruction-cellulases for which there are a large number of gene sequences available, xylanases and ligninases etc. for which there is scant gene and enzyme data available in the database.